CBIIT TechScouts

'Inspiring innovation through forums for idea-sharing'

CBIIT TechScouts is a forum for promoting continuous improvement across CBIIT through the cross-fertilization of ideas, experiences and recommendations. This forum is designed to foster new collaborations, learn about opportunities to better serve our customers, engage CBIIT more broadly, and raise awareness of new techniques and technologies that promote innovation. Central to the CBIIT TechScouts is the focus on gathering insight from the community with ideas and experiences on how emerging information can be used to improve scientific productivity and accelerate cancer research.

Table of Contents

- Table of Contents
- Topics Archive
- Show and Tell Archive
- Attachments

Topics Archive

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic (CLICK FOR MORE INFO)</th>
<th>Author</th>
<th>Topic ID</th>
<th>Summary</th>
</tr>
</thead>
</table>
| 10 Oct   | Arlington VA AWS Meetup list: "Call for presenters" | Sean Davis   | TS-0057  | There might be room for CBIIT to participate? I think that other government agencies would be very interested in hearing how NCI is approaching AWS.

Natasha Clark (Co-Organizer) sent a message to the Arlington VA AWS Meetup mailing list - call for presenters

Hi Everyone!

I wanted to reach out to see if any of you would be interested in presenting at our October session, scheduled for Thursday, October 26th at Excella.

Depending on interest, we can either make this a lightning talk session or keynote depending on length. As always, we are always interested in all things AWS and below are a few topics we have had interest for in the past.

If any of these sound like they might be right up your alley, or if you have another topic in mind, please reach out to me via the messaging feature on the meetup.com site and I can help coordinate.

Looking forward to seeing you all later this month!

Natasha

- Intro into lambda/ serverless
- Cost effective AWS practices
- All things migration
- Container management

| 04 Oct   | The big hack | Richard Finney | TS-0056  | Bloomberg is headlining a hardware hack: The Big Hack: How China Used a Tiny Chip to Infiltrate U.S. Companies


Nested on the servers’ motherboards, the testers found a tiny microchip, not much bigger than a grain of rice, that wasn’t part of the boards’ original design. Amazon reported the discovery to U.S. authorities, sending a shudder through the intelligence community. ….

The companies’ denials are countered by six current and former senior national security officials, who — … — detailed the discovery of the chips and the government’s investigation.

BUT … Amazon and Apple are denying they’ve been compromised …


Amazon.com Inc. (AMZN) and Apple Inc. (AAPL) have denied claims that a secret microchip was found embedded in servers linked to Elemental Technologies, a video compressing service purchased by Amazon in 2015, amid concerns that government hackers in China were able to infiltrate U.S. corporate data.

Bloomberg reported Thursday that the chip, found on a server made by San Jose, Calif.-based Super Micro Computer Inc (SMCI) via subcontractors in China through a contract with Elemental, could be used to infiltrate a host of computer networks linked to both major U.S. companies as well as portions of the U.S. government’s national security system.
<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
<th>Speaker</th>
<th>TS</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>If you're interested in learning more about blockchain technology and its potential applicability in federal work, be aware that NIST just released NISTIR 8202. This “Blockchain Technology Overview” is a technical publication that examines the history, scope, and other characteristics of blockchain technology. NISTIR 8202 discusses various blockchain implementation approaches, existing limitations and misconceptions surrounding blockchain, and several areas of consideration for federal agencies and organizations seeking to understand and manage blockchain technology. It is also an introductory document meant to provide the foundation for a planned series of publications on more specific aspects of blockchain.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26 Sep</td>
<td>Software disenchantment - Everything is unbearably slow</td>
<td>Richard Pinney</td>
<td>TS-0054</td>
<td><a href="http://tonsky.me/blog/disenchantment/">http://tonsky.me/blog/disenchantment/</a></td>
</tr>
<tr>
<td></td>
<td>A lament on the current state of software: Modern text editors have higher latency than 42-year-old Emacs. Text editors! What can be simpler? On each keystroke, all you have to do is update tiny rectangular region and modern text editors can’t do that in 16ms. It’s a lot of time. A LOT. A 3D game can fill the whole screen with hundreds of thousands (10’s) of polygons in the same 16ms and also process input, recalculate the world and dynamically load/unload resources. How come?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>With the release of the gen 4 Apple Watch, which includes (or will include) an ECG and AFib monitoring features, we will probably be seeing a lot more stories like this.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 Sep</td>
<td>Local news: Montgomery County Hearing on ZTA 18-11</td>
<td>Robert Wynne</td>
<td>TS-0552</td>
<td>(The following public information may be of interest to anyone living and/or working in Montgomery County.)</td>
</tr>
<tr>
<td></td>
<td>Wireless Facilities Hearing on ZTA 18-11 and Map</td>
<td></td>
<td></td>
<td>A public hearing will be held regarding ZTA 18-11 on Sept. 25 at 7:30 p.m. in the third-floor hearing room of the Council Office Building at 100 Maryland Avenue in Rockville.</td>
</tr>
<tr>
<td></td>
<td>The Wireless Facilities Map describing new towers and 5G mini-towers, as proposed to the County Council, is publicly available. There are multiple locations planned for 5G mini-towers and in many communities across Montgomery County more than 30’ from homes, as well as new mobile utility poles. <a href="https://gis.montgomerycountymd.gov/Wireless/">https://gis.montgomerycountymd.gov/Wireless/</a></td>
<td></td>
<td></td>
<td>Resources</td>
</tr>
<tr>
<td></td>
<td>With higher assurance levels for protected data, subscribers can easily manage this data and share it securely and appropriately with collaborators. These new features especially benefit organizations and projects where protected data is shared by multiple researchers, such as institutions with secure data enclaves, multi-institutional studies using clinical data, and facilities distributing sensitive data to investigators and their collaborators.</td>
<td></td>
<td></td>
<td>Read the announcement to get more details, or register for a live Q&amp;A webinar on October 24.</td>
</tr>
<tr>
<td>12 Sep</td>
<td>Announcing Globus. Support for Protected Data</td>
<td>Sean Davis</td>
<td>TS-0051</td>
<td>We're excited to announce availability of new Globus features for managing protected data, including HIPAA-regulated data and personally identifiable information.</td>
</tr>
<tr>
<td></td>
<td>With higher assurance levels for protected data, subscribers can easily manage this data and share it securely and appropriately with collaborators. These new features especially benefit organizations and projects where protected data is shared by multiple researchers, such as institutions with secure data enclaves, multi-institutional studies using clinical data, and facilities distributing sensitive data to investigators and their collaborators.</td>
<td></td>
<td></td>
<td>Read the announcement to get more details, or register for a live Q&amp;A webinar on October 24.</td>
</tr>
<tr>
<td>04 Sep</td>
<td>Lecture Announcement: Containerization for Reproducible Bioinformatics Research</td>
<td>Sean Davis</td>
<td>TS-0050</td>
<td>Containerization for Reproducible Bioinformatics Research</td>
</tr>
<tr>
<td></td>
<td>Date: Tuesday, September 4, 2018</td>
<td></td>
<td></td>
<td>Time: 11:00 am – 12:00 pm</td>
</tr>
<tr>
<td></td>
<td>Location: Room E1/E2, Natcher Conference Center (NIH Building 45)</td>
<td></td>
<td></td>
<td>Registration: No pre-registration is required. Seating is first come first serve.</td>
</tr>
<tr>
<td></td>
<td>As computational work becomes increasingly embedded in biomedical research practices, computational reproducibility has become an issue of increasing importance. Computational reproducibility requires that other researchers are able to deploy and use software and analysis workflows in their own computing environments. Platforms like Docker and Singularity allow the creation and configuration of software containers, which can be distributed and deployed across a range of systems. This lecture, presented by Steve Tsang, will give an introductory overview of containerization and how containers can facilitate reproducible bioinformatics research, providing examples from the NCI Cloud Resources and various hackathons.</td>
<td></td>
<td></td>
<td>Seating is limited, but the presentation will be available through Webex (calendar invite attached).</td>
</tr>
<tr>
<td></td>
<td><a href="https://martinfowler.com/articles/agile-aus-2018.html">https://martinfowler.com/articles/agile-aus-2018.html</a></td>
<td></td>
<td></td>
<td>This article was included in the most recent issue of NIH’s IRP Weekly: Doctor Data: How Computers Are Invading the Clinic / AI for Biomedical Research</td>
</tr>
<tr>
<td></td>
<td>Doctor Data: How Computers Are Invading the Clinic / AI for Biomedical Research</td>
<td>Carl McCabe</td>
<td>TS-0048</td>
<td>This article was included in the most recent issue of NIH’s IRP Weekly: Doctor Data: How Computers Are Invading the Clinic / AI for Biomedical Research</td>
</tr>
<tr>
<td></td>
<td><a href="https://irp.nih.gov/blog/post/2018/08/doctor-data-how-computers-are-invading-the-clinic">https://irp.nih.gov/blog/post/2018/08/doctor-data-how-computers-are-invading-the-clinic</a></td>
<td></td>
<td></td>
<td>This article was included in the most recent issue of NIH’s IRP Weekly: Doctor Data: How Computers Are Invading the Clinic / AI for Biomedical Research</td>
</tr>
</tbody>
</table>
We (Bioconductor) have created an online and published set of workshop resources that we used for our annual conference. We did so using the open source Bookdown package (https://bookdown.org/yihui/bookdown/) in a collaborative editing effort that resulted in 388 pages from 19 contributors in just over 8 weeks. There is an associated Amazon Machine Image that was used to test build the materials and then each conference participant received his/her own instance for the duration of the conference.

Materials are here:
https://bioconductor.github.io/BiocWorkshops/ (html)

Feel free to contact me to discuss the process in more detail.

Three things in computers are hard: cache invalidation and off by one errors.

"That impact is now a little larger. Researchers from Graz University of Technology, including one of the original Meltdown discoverers, Daniel Gruss, have described NetSpectre: a fully remote attack based on Spectre. With NetSpectre, an attacker can remotely read the memory of a victim system without running any code on that system": (from ars technica)

News:
https://duckduckgo.com/?q=netspectre+&t=ffsb&iar=news&ia=news

original reporting paper:
https://misc0110.net/web/files/netspectre.pdf
ouch.

A one size fits all database doesn't fit anyone

New, registration is open for the next edition of DigitalGov University Emerging Technology Leadership Series, a new pilot to enhance the modern federal workforce with training, education, and awareness of emerging technologies including Artificial Intelligence, Robotic Process Automation, Blockchain, Social Technologies, and Virtual/Augmented Reality.

Packer (https://packer.io) is a toolkit that implements "infrastructure-as-code" (https://en.wikipedia.org/wiki/Infrastructure_as_Code) for building Amazon Machine Images (AMIs). At the annual Bioconductor conference, 50% of the conference is devoted to hands-on coding and we supply every participant with a custom AMI. To ensure reproducibility and reusability of the AMI, we used packer to automate the creation of the AMI directly from a json file, eliminating any hand-editing or configuration of the machine. I describe this process very briefly in a blog post here:
I thought it might be useful for a few folks.

This might be of interest to a few folks. A one size fits all database doesn't fit anyone.

We'd like to invite you to a private delivery of our new Systems Operations on AWS course, which will be publicly released later this year. This instructor-led preview is scheduled for 7 – 9 August in Herndon, VA. We’re offering seats at 50% off our standard price and asking for detailed feedback on class content and delivery.

In Systems Operations on AWS, we teach individuals how to create automatable and repeatable deployments of networks and systems on the AWS platform. We will explore the AWS features and tools related to configuration and deployment and best practices for configuring and deploying systems. You will also learn how to:

- Use standard AWS infrastructure features such as Amazon Virtual Private Cloud (VPC), Amazon Elastic Compute Cloud (EC2), Elastic Load Balancing (ELB), and AWS Auto Scaling from the command line
- Use AWS CloudFormation and other automation technologies to produce stacks of AWS resources
- Build virtual private networks with Amazon VPC

Seats for this invitation-only preview are $900 (a 50% savings on the full price). Space is limited, so we encourage you to register today.

Microsoft is Said to Have Agreed to Acquire Coding Site GitHub

"Data infuses intelligence in to every business."

Perhaps a group or two will be interested in implementing. Note that CBIT the NCI Cleversafe object storage system could be a nice technological base for a data lake.
<table>
<thead>
<tr>
<th>Date</th>
<th>Event Description</th>
<th>Author</th>
<th>Event ID</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>07 May 2018</td>
<td>Biowulf transitioning from RHEL6/CentOS6 to RHEL7/CentOS7 in June</td>
<td>Sean Davis</td>
<td>TS-0036</td>
<td>If you are working with Biowulf users (and NCI has MANY), this may be of interest to you. If there are CBIIT groups who would like to learn more about Biowulf, let me know and we can probably organize either a tour of Biowulf facilities or a Biowulf staff visit. For general information about Biowulf, the NIH enterprise HPC system, see: <a href="https://hpc.nih.gov/">https://hpc.nih.gov/</a></td>
</tr>
<tr>
<td>06 May 2018</td>
<td>Blockchain explained in 7 python functions</td>
<td>Sean Davis</td>
<td>TS-0035</td>
<td>Ever wonder about what blockchain actually IS?</td>
</tr>
<tr>
<td>03 May 2018</td>
<td>Python pip support ending for TLS versions less than 1.2</td>
<td>Rohit Paul</td>
<td>TS-0034</td>
<td>I was just bit by this on my Mac, so passing along in case anyone experiences (or has experienced) being unable to install/upgrade packages using 'pip' due to a TLS-related error. <a href="http://pyfound.blogspot.ca/2017/01/time-to-upgrade-your-python-tls-v12.html">http://pyfound.blogspot.ca/2017/01/time-to-upgrade-your-python-tls-v12.html</a> This affected my El Capitan 10.11 system; the default Python on macOS 10.12+ is hopefully unaffected. Installing the latest Python 2.x using Homebrew worked fine for me, after upgrading Ruby as well. Fun stuff.</td>
</tr>
<tr>
<td>11 Apr 2018</td>
<td>NCI Containers and Workflow Seminar</td>
<td>Carl McCabe</td>
<td>TS-0033</td>
<td>Here's a talk to the NIH Data Science SIG that several people across CBIIT may be interested in. NCI Containers and Workflows Interest Group / NIH Data Science Lectures Joint Seminar  Bio-Docklets: virtualization containers for single-step execution of NGS pipelines  Presenter: Konstantinos Krampis - Associate Professor, Hunter College, City University of New York; Faculty, Weill Cornell Medical College  WebEx URL: <a href="https://cbiit.webex.com/cbiit/j.php?MTID=ma310d94955bd365d114c5a469359aca">https://cbiit.webex.com/cbiit/j.php?MTID=ma310d94955bd365d114c5a469359aca</a>  Meeting number (access code): 731 927 985</td>
</tr>
<tr>
<td>07 Mar 2018</td>
<td>Debian GNU/Linux for WSL</td>
<td>Carl McCabe</td>
<td>TS-0032</td>
<td>Debian GNU/Linux for WSL now available in the Windows Store</td>
</tr>
</tbody>
</table>

This is an update of a previous announcement about Linux on Linux on WSL: New distros coming to Bash/WSL via Windows Store https://blogs.msdn.microsoft.com/commandline/2017/06/11/new-distros-coming-to-bashws/
The White House’s American Technology Council and Office of American Innovation on modernizing federal technology published specific recommendations to jumpstart a new wave of modernization efforts by accelerating cloud adoption, consolidating networks and prioritizing key applications for needed upgrades.

Now the daunting task of implementing these recommendations sits within agencies, and it is not a one size fits all proposition. Where do they begin to successfully move away from expensive legacy infrastructure? How do they transition to a more secure, agile, and cost-effective technology ecosystem, much of which will be supported by shared services?

Join i360Gov and senior level technology leaders from government and industry as we provide an overview of current initiatives and solutions to the many IT modernization challenges agencies face, such as:

- Maintenance that often requires immediate attention and runs the risk of breaking integrations and upgrades
- Legacy solutions that are unable to properly communicate between on-premises, mobile, and the cloud
- Citizen facing services not designed for today’s technology environment

You will also learn about Identity, the hidden accelerator to IT modernization, and how by creating a single solution for identity, agencies can speed up digital and cloud programs that will enable you to:

- Reduce costs and architecture complexity
- Securely connect any employee, vendor, partner or citizen to any resource, on premise or in the cloud
- Make administrators self-sufficient and decrease reliance on customization
- Scale seamlessly as you move services into the cloud

**Webinar Presenters**

- Dr. Ronald Ross, Computer Scientist, NIST Fellow
- David Hogue, Technical Director, Cybersecurity Threat Operations Center, NSA
- Joe Diamond, Director, Cybersecurity Strategy, Okta

Register now for this complimentary, educational webinar. As long as you register now, you will receive the link. If you are unable to attend the live webinar, you can use the link to watch the recording at your convenience.
Interested in Agile? Keen on DevOps?

Please join the Advanced Technology Academic Research Center next Thursday, March 1 for the premier government Agile and DevOps event of the year, the ATARC Federal DevOps Summit at the Marriott Metro Center in Washington, D.C. This educational symposium for Federal IT practitioners is free to government and eligible for 7.5 CPE credits. To view the agenda and register, please visit: www.fedsummits.com/devops/

A Visionary Panel on DevOps and Government Transformation will include Cris Brown, Master Data Management Program Manager, NRC; Peter Burkholder, Innovation Specialist, GSA 18F; Jennifer Hosvar, Digital Services Expert, DHS; and Evan Lee, Chief Technology Officer, HHS OIG. Tom Temmin of Federal News Radio will serve as moderator.

An all-star list of Federal IT thought leaders include: David Lamirande, Chief Technology Officer, DHS ICE; Simmons Lough, Software Architect, USPTO; Navin Verma, Chief Technology Officer, GSA; and Hasan Yasar, Technical Manager & Adjunct Faculty Member, Software Engineering Institute, Carnegie Mellon University.

The afternoon will feature the MITRE-ATARC DevOps Collaboration Symposium with government, academic and private industry SMEs who will brainstorm and whiteboard during five concurrent sessions on: SecDevOps; DevOps Implementation with Cloud; DevOps Culture; DevOps Testing; and DevOps in Health IT.

Session topics:
1. SecDevOps
2. DevOps Implementation with Cloud
3. DevOps Culture
4. DevOps Testing
5. DevOps in Health IT

In the February 9th issue of the NIH Record:

NIH Supercomputers Have Come a Long Way

BY DANA TALEN


I don’t know how long this deal will last, but Amazon currently (Friday 5pm) has Pro Git 2, by Scott Chacon and Ben Staub, on sale at a very affordable price of $0.00. It is the Kindle edition, not a paper copy. If you’re interested, here’s the link:

https://www.amazon.com/gp/product/B01ISNIKES

Melina Scotto: ‘A reminder that the hash used in GIT is SHA1. Deprecated for government use since 2013. Great to play with, though.’

Jeff Shilling: ‘Since Git is not using SHA1 for cryptographic functions but as a hash of the files to determine if changes to the files have taken place, it doesn’t pose a security concern.’

This is not something we can use right now (and maybe not ever in the Federal govt), but it is an interesting peak into the future.

Golem
https://golem.network/

“Golem is a global, open source, decentralized supercomputer that anyone can access. It is made up of the combined power of users’ machines, from PCs to entire data centers. Golem is capable of computing a wide variety of tasks, from CGI rendering, through machine learning to scientific computing. Golem’s limitations are only defined by software developers’ creativity. Golem creates a decentralized sharing economy of computing power and supplies software developers with a flexible, reliable and cheap source of computing power.”

Golem does have competitors, e.g. https://iex.ec/

NCI has established the Genomic Data Commons (GDC, https://gdc.cancer.gov) as a home for NCI cancer genomics datasets. One feature of the GDC is that it uses UUIDs for “everything”. However, most cancer researchers think in terms of legacy “barcodes” when working with these datasets. Given the increasing importance of web-based APIs and their use for NCI data, I wrote a quick blog post using an R client for these data and, specifically, for translating from UUID back to “barcodes”:


Feel free to pass along.

My session on truly introductory R is from 10-12 today. Materials (we won’t cover all of them) are here:

https://seandavi.github.io/ITR/
04 Dec 2017 Biocountucto r, Software for Genomic Data Science Sean Davis TS-0023 I am giving a talk at 11am about Biocountuctor, a large software project partially funded by NCI and the CBIIIT ITCR program. The talk will be the TE406 at 11am and is only 20 minutes long—part of a morning-long conference open to anyone sponsored by the Center of Excellence in Integrative Cancer Biology and Genomics.

Slides are posted here:
http://bit.ly/2bKgJG

Here is the abstract:
Progress in biotechnology is continually leading to new types of data, resulting in data sets that are rapidly increasing in volume, resolution and diversity. The promise of unprecedented advances in our understanding of biological systems and in medicine is challenged by complexity and volume of data also challenge scientists’ ability to analyze them. Meeting this challenge requires continuous improvements in analytical methods and capable, usable software tools implementing them. Biocountuctor is a well-established open-source, open-development software project for the analysis and comprehension of high-throughput data in genomics and molecular biology. The project aims to enable interdisciplinary research, collaboration and rapid development of scientific software. Based on the statistical programming language R, Biocountuctor comprises 1473 interopenable packages contributed by a large, diverse community of scientists. These packages undergo formal initial review and continuous automated testing. Each package includes documentation and working example use cases. Biocountuctor supports many types of high-throughput sequencing data (including DNA, RNA, chromatin immunoprecipitation, Hi-C, metagenomes and ribosome profiling) and associated annotation resources; contains mature facilities for microarray analysis; and covers proteomic, metabolic, flow cytometry, quantitative imaging, chemoinformatic and other high-throughput data. Biocountuctor package interoperability enables the rapid creation of workflows combining and integrating multiple data types and tools for statistical inference, regression, network analysis, machine learning and visualization at all stages of a project from data generation to publication. A large and growing community of researchers and users contribute to ongoing development, online support, and education. The influence of the project is evidenced by more than 250,000 downloads per year and tens of thousands of citations in the literature. I will present an overview of the project for prospective users and contributors.

02 Dec 2017 Protect Against Secrets in Git Repositories Sean Davis TS-0022 I wrote a blog post about an embarrassing but educational experience including the solution I implemented for myself to keep it from happening again. Perhaps someone else will find my experience useful.

http://bit.ly/29tco0k

17 Nov 2017 GitHub - Security Alerts and Dependency Graphs Carl McCabe TS-0021 GitHub recently rolled out a new feature allowing you to see a graph of your project's dependencies. And yesterday they took this a step further by providing you with alerts when vulnerabilities are detected in any of those dependencies. Dependency Graphs are enabled automatically for public repos, but must be enabled optionally for private repos. Currently, this works for Ruby and JavaScript, but Python is coming sometime in 2018.

Dependency Graphs
https://help.github.com/articles/listing-the-packages-that-a-repository-depends-on/

Security Alerts

16 Nov 2017 Agricultural Data Ecosystem Talk Sean Davis TS-0020 I recently visited the USDA to give a talk with thoughts on an Agricultural Data Ecosystem, inspired in part by the work done by pioneering source technologies and AWS services. In the future, we'll be moving Nephele components to the custom PaaS we created called Monarch and hosting it under our architecture.

For background, Nephele, discussed in the presentation, is a project that a team in BCBB has been working on for several years and provided some guidance for best practices and improvements now that we are launching a custom platform as a service using open source technologies and AWS services. In data science, we are moving Nephele components to the custom PaaS we created called Monarch and hosting it under our architecture.

https://www.youtube.com/watch?v=rLkBwWv0Hdc&list=PLhr1KZpdzukePsKIUofhgp50b63-5yr1V&index=66

20 Sep 2017 Bored With Your Fitbit? These Cancer Researchers Aren’t (Wired.com)

BORED WITH YOUR FITBIT? THESE CANCER RESEARCHERS AREN’T

https://www.wired.com/story/bored-with-your-fitbit-these-cancer-researchers-arent/

11 Aug 2017Malware Encoded Into DNA Can Hack the Computer That Reads It
Joel Yabouet TS-0018 What if someone stores a malicious program into DNA, just like an infected USB storage, to hijack the computer that reads it?

A team of researchers from the University of Washington in Seattle have demonstrated the first successful DNA-based exploit of a computer system that executes the malicious code written into the synthesized DNA strands while reading it.

Here is the white paper :

02 Aug 2017 DigiCert is Acquiring Symantec’s Website Security Business
Jesse Bocinski TS-0017 DigiCert is acquiring Symantec’s web security business, a consolidation of the public commercial certificate market.

http://bit.ly/2AUYy2w

06 Jul 2017 Pretty amazing feats of science and engineering to get us to 2020.
Warren Kibbe TS-0016

https://en.wikipedia.org/wiki/5纳米ometer

For comparison, a single benzene ring is about 500 picometers, or 0.5 nanometers

29 Jun 2017
Monarch Made Twitter
Amy Ginzelt TS-0015

I wanted to share a session that was presented at the AWS Public Sector Summit by NIAID that discussed Next-Generation Medical Analysis leveraging the cloud. Our CSRA team at NIAID assisted by developing a Paas solution they’ve named Monarch, discussed in the presentation, that assists with this work leveraging a full DevOps pipeline.

For background, Nephele, discussed in the presentation, is a project that a team in BCBB has been working on for several years now. Due to the infrastructure group (OEB) not having a defined service offering for public cloud at NIAID years ago when Nephele started, the project team went on their own and architected a solution using AWS. Within the last year, CSRA consulted with them and provided some guidance on how to develop with the public cloud services and improvements now that we are launching a custom platform as a service using open source tools and AWS services. In the future, we’ll be moving Nephele components to the custom Paas we created called Monarch and hosting it under our architecture.

https://www.youtube.com/watch?v=UAUwHVvOHzo&feature=plcp&list=PLhr1KZpdzukePsKIUofhgp50b63-5yr1V&index=66

29 Jun 2017 DCEG Linkage: 3D Printing in Radiation Research
Carl McCabe TS-0014

Here’s a unique example of 3D printing right on site in the NCI Shady Grove building. This is from Choosnik Leel’s group in DCEG’s Radiation Epidemiology Branch.


16 Jun 2017 SciCon: China Shatters Spooky Action at a Distance Record, Prep for Quantum Internet
Carl McCabe TS-0013

Here’s more on the long-term future (ignore the immediate or mid-term geopolitical implications).

With Maru, your phone is your PC——you connect it to a keyboard and monitor whenever you need the desktop environment. Obviously this isn’t ready for government use yet, but it illustrates the path toward convergence into single device personal computing.

Great piece in the April 3 New Yorker on medical diagnostics and machine learning. They used Thrun and Hinton’s Columbia work with TS-0003

The Bioconductor project

Interesting article about the history—and potential future—of Windows: http://www.newyorker.com/magazine/2017/04/03/ai-versus-md

For anyone paying attention to technology trends in statistical analysis and data science, here’s a report from an R-focused data science scientist. Carl McCabe

If you have questions, feel free to contact me or the Biowulf staff.

Globus, a cloud-authenticated data management and transfer platform, will be hosting a user-focused webinar on Tuesday, May 16th, to benefit those interested in exchanging datasets across a variety of sources. The webinar will provide a high-level overview of Globus, steps to start using the service, and common use cases along the following topics:

1. When, where, and why to use Globus?
2. NIH account specifics—distinction from Globus Plus
3. What do system administrators need to set up managed endpoints?
4. Which endpoints are already set up?
5. How to set up Globus on your own desktop
6. How to transfer and share data
7. If sharing with collaborator, what info does collaborator need?
8. What do you need to give to collaborator?
9. New Globus command line interface, allowing users to script their transfers
10. Encryption, verification, and expected data transfer speeds as compared to other resources (e.g. FTP)

Great piece in the April 3 New Yorker on medical diagnostics and machine learning. They used Thrun and Hinton’s Columbia work with TS-0003

The Bioconductor project

Interesting article about the history—and potential future—of Windows: http://www.newyorker.com/magazine/2017/04/03/ai-versus-md

For anyone paying attention to technology trends in statistical analysis and data science, here’s a report from an R-focused data science scientist. Carl McCabe

If you have questions, feel free to contact me or the Biowulf staff.

Globus, a cloud-authenticated data management and transfer platform, will be hosting a user-focused webinar on Tuesday, May 16th, to benefit those interested in exchanging datasets across a variety of sources. The webinar will provide a high-level overview of Globus, steps to start using the service, and common use cases along the following topics:

1. When, where, and why to use Globus?
2. NIH account specifics—distinction from Globus Plus
3. What do system administrators need to set up managed endpoints?
4. Which endpoints are already set up?
5. How to set up Globus on your own desktop
6. How to transfer and share data
7. If sharing with collaborator, what info does collaborator need?
8. What do you need to give to collaborator?
9. New Globus command line interface, allowing users to script their transfers
10. Encryption, verification, and expected data transfer speeds as compared to other resources (e.g. FTP)

Great piece in the April 3 New Yorker on medical diagnostics and machine learning. They used Thrun and Hinton’s Columbia work with TS-0003

The Bioconductor project

Interesting article about the history—and potential future—of Windows: http://www.newyorker.com/magazine/2017/04/03/ai-versus-md

For anyone paying attention to technology trends in statistical analysis and data science, here’s a report from an R-focused data science scientist. Carl McCabe

If you have questions, feel free to contact me or the Biowulf staff.

Globus, a cloud-authenticated data management and transfer platform, will be hosting a user-focused webinar on Tuesday, May 16th, to benefit those interested in exchanging datasets across a variety of sources. The webinar will provide a high-level overview of Globus, steps to start using the service, and common use cases along the following topics:

1. When, where, and why to use Globus?
2. NIH account specifics—distinction from Globus Plus
3. What do system administrators need to set up managed endpoints?
4. Which endpoints are already set up?
5. How to set up Globus on your own desktop
6. How to transfer and share data
7. If sharing with collaborator, what info does collaborator need?
8. What do you need to give to collaborator?
9. New Globus command line interface, allowing users to script their transfers
10. Encryption, verification, and expected data transfer speeds as compared to other resources (e.g. FTP)

Great piece in the April 3 New Yorker on medical diagnostics and machine learning. They used Thrun and Hinton’s Columbia work with TS-0003

The Bioconductor project

Interesting article about the history—and potential future—of Windows: http://www.newyorker.com/magazine/2017/04/03/ai-versus-md

For anyone paying attention to technology trends in statistical analysis and data science, here’s a report from an R-focused data science scientist. Carl McCabe

If you have questions, feel free to contact me or the Biowulf staff.

Globus, a cloud-authenticated data management and transfer platform, will be hosting a user-focused webinar on Tuesday, May 16th, to benefit those interested in exchanging datasets across a variety of sources. The webinar will provide a high-level overview of Globus, steps to start using the service, and common use cases along the following topics:

1. When, where, and why to use Globus?
2. NIH account specifics—distinction from Globus Plus
3. What do system administrators need to set up managed endpoints?
4. Which endpoints are already set up?
5. How to set up Globus on your own desktop
6. How to transfer and share data
7. If sharing with collaborator, what info does collaborator need?
8. What do you need to give to collaborator?
9. New Globus command line interface, allowing users to script their transfers
10. Encryption, verification, and expected data transfer speeds as compared to other resources (e.g. FTP)

Great piece in the April 3 New Yorker on medical diagnostics and machine learning. They used Thrun and Hinton’s Columbia work with TS-0003

The Bioconductor project

Interesting article about the history—and potential future—of Windows: http://www.newyorker.com/magazine/2017/04/03/ai-versus-md

For anyone paying attention to technology trends in statistical analysis and data science, here’s a report from an R-focused data science scientist. Carl McCabe

If you have questions, feel free to contact me or the Biowulf staff.

Globus, a cloud-authenticated data management and transfer platform, will be hosting a user-focused webinar on Tuesday, May 16th, to benefit those interested in exchanging datasets across a variety of sources. The webinar will provide a high-level overview of Globus, steps to start using the service, and common use cases along the following topics:

1. When, where, and why to use Globus?
2. NIH account specifics—distinction from Globus Plus
3. What do system administrators need to set up managed endpoints?
4. Which endpoints are already set up?
5. How to set up Globus on your own desktop
6. How to transfer and share data
7. If sharing with collaborator, what info does collaborator need?
8. What do you need to give to collaborator?
9. New Globus command line interface, allowing users to script their transfers
10. Encryption, verification, and expected data transfer speeds as compared to other resources (e.g. FTP)

Great piece in the April 3 New Yorker on medical diagnostics and machine learning. They used Thrun and Hinton’s Columbia work with TS-0003

The Bioconductor project

Interesting article about the history—and potential future—of Windows: http://www.newyorker.com/magazine/2017/04/03/ai-versus-md

For anyone paying attention to technology trends in statistical analysis and data science, here’s a report from an R-focused data science scientist. Carl McCabe

If you have questions, feel free to contact me or the Biowulf staff.

Globus, a cloud-authenticated data management and transfer platform, will be hosting a user-focused webinar on Tuesday, May 16th, to benefit those interested in exchanging datasets across a variety of sources. The webinar will provide a high-level overview of Globus, steps to start using the service, and common use cases along the following topics:

1. When, where, and why to use Globus?
2. NIH account specifics—distinction from Globus Plus
3. What do system administrators need to set up managed endpoints?
4. Which endpoints are already set up?
5. How to set up Globus on your own desktop
6. How to transfer and share data
7. If sharing with collaborator, what info does collaborator need?
8. What do you need to give to collaborator?
9. New Globus command line interface, allowing users to script their transfers
10. Encryption, verification, and expected data transfer speeds as compared to other resources (e.g. FTP)

Great piece in the April 3 New Yorker on medical diagnostics and machine learning. They used Thrun and Hinton’s Columbia work with TS-0003

The Bioconductor project

Interesting article about the history—and potential future—of Windows: http://www.newyorker.com/magazine/2017/04/03/ai-versus-md

For anyone paying attention to technology trends in statistical analysis and data science, here’s a report from an R-focused data science scientist. Carl McCabe

If you have questions, feel free to contact me or the Biowulf staff.

Globus, a cloud-authenticated data management and transfer platform, will be hosting a user-focused webinar on Tuesday, May 16th, to benefit those interested in exchanging datasets across a variety of sources. The webinar will provide a high-level overview of Globus, steps to start using the service, and common use cases along the following topics:

1. When, where, and why to use Globus?
2. NIH account specifics—distinction from Globus Plus
3. What do system administrators need to set up managed endpoints?
4. Which endpoints are already set up?
5. How to set up Globus on your own desktop
6. How to transfer and share data
7. If sharing with collaborator, what info does collaborator need?
8. What do you need to give to collaborator?
9. New Globus command line interface, allowing users to script their transfers
10. Encryption, verification, and expected data transfer speeds as compared to other resources (e.g. FTP)

Great piece in the April 3 New Yorker on medical diagnostics and machine learning. They used Thrun and Hinton’s Columbia work with TS-0003

The Bioconductor project

Interesting article about the history—and potential future—of Windows: http://www.newyorker.com/magazine/2017/04/03/ai-versus-md

For anyone paying attention to technology trends in statistical analysis and data science, here’s a report from an R-focused data science scientist. Carl McCabe

If you have questions, feel free to contact me or the Biowulf staff.

Globus, a cloud-authenticated data management and transfer platform, will be hosting a user-focused webinar on Tuesday, May 16th, to benefit those interested in exchanging datasets across a variety of sources. The webinar will provide a high-level overview of Globus, steps to start using the service, and common use cases along the following topics:

1. When, where, and why to use Globus?
2. NIH account specifics—distinction from Globus Plus
3. What do system administrators need to set up managed endpoints?
4. Which endpoints are already set up?
5. How to set up Globus on your own desktop
6. How to transfer and share data
7. If sharing with collaborator, what info does collaborator need?
8. What do you need to give to collaborator?
9. New Globus command line interface, allowing users to script their transfers
10. Encryption, verification, and expected data transfer speeds as compared to other resources (e.g. FTP)
Monitoring changes to infrastructure and data is a very important part of a robust and secure infrastructure. This service allows ongoing logging of API calls (so, changes) to infrastructure and even supports S3 events (for changes to data).

https://aws.amazon.com/cloudtrail/

Other cloud providers offer similar services directly or via third-parties.

---

Show and Tell Archive

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Who</th>
<th>Presentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>09 Oct 2018</td>
<td>Introduction to Blockchain</td>
<td>Manish Malhotra - Chairman &amp; CEO, Unnissant, Inc.</td>
<td></td>
</tr>
</tbody>
</table>

---

Attachments

<table>
<thead>
<tr>
<th>File</th>
<th>Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microsoft Powerpoint Presentation CBIIT Sync Meeting_Globus_MK.pptx</td>
<td>Apr 11, 2017 by Kimbrough, Miles (NIH/NCI) [C]</td>
</tr>
<tr>
<td>PDF File Introduction to Blockchain.pdf</td>
<td>Sep 24, 2018 by Kimbrough, Miles (NIH/NCI) [C]</td>
</tr>
<tr>
<td>PNG File HPC-Blockchain-intro-image.png</td>
<td>Dec 03, 2018 by Frost, Ruth (NIH/NCI) [C]</td>
</tr>
</tbody>
</table>

Download All

---

Welcome to CBIIT TechScouts!

We are a forum for promoting continuous improvement across CBIIT through the cross-fertilization of ideas, experiences and recommendations.

The goal of CBIIT TechScouts is to collectively improve scientific productivity through the sharing of knowledge.

What We Do

CBIIT TechScouts is designed to foster new collaborations, learn about opportunities to better serve our customers, engage CBIIT more broadly, and raise awareness of new techniques and technologies that promote innovation.
Central to the CBIIT Tech Scouts is the focus on gathering insight from the community with ideas and experiences on how emerging information, data-oriented, and software technologies can be used to improve scientific productivity and accelerate cancer research.

Our Service Includes:

1. **An email distribution list**, allowing members to stay updated on the latest topics and trends
   a. To join: email Miles Kimbrough, subject 'TechScouts Access Request'

2. **An annotated archive of topics and presentations**, organized for future reference

3. **A monthly summary found on CBIIT Central**, providing a lightweight overview of recent trends

4. **And finally, Monthly Show and Tell Sessions**, opening the door for new ideas to be presented both in person and virtually
   a. To present: email Miles Kimbrough with topic and availability

**Questions?**

**General Support**: Miles Kimbrough | miles.kimbrough@nih.gov | 240.276.5251

**Consultation and Guidance**: Eric Stahlberg | eric.stahlberg@nih.gov | 240.276.6729

**Technical Support**: George Zaki | george.zaki@nih.gov | 240.276.5171